



HAL
open science

Composite Visual Mapping for Time Series Visualization

Ali Jabbari, Renaud Blanch, Sophie Dupuy-Chessa

► **To cite this version:**

Ali Jabbari, Renaud Blanch, Sophie Dupuy-Chessa. Composite Visual Mapping for Time Series Visualization. Proceedings of the 11th IEEE Pacific Visualization Symposium (PacificVis 2018), Apr 2018, Kobe, Japan. pp.116-124, <10.1109/PacificVis.2018.00023>. <hal-03223754>

HAL Id: hal-03223754

<https://hal.science/hal-03223754v1>

Submitted on 11 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Composite Visual Mapping for Time Series Visualization

Ali Jabbari*

Renaud Blanch†

Sophie Dupuy-Chessa‡

Univ. Grenoble-Alpes, LIG
F-38000 Grenoble, France

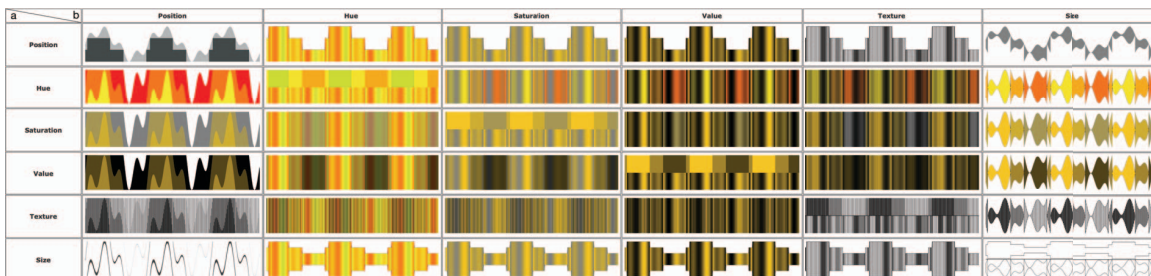


Figure 1: A systematic overview of possible combination of visual channels (here, using modulo transformation: $v = a \times M + b$). Each sub-attribute a and b is mapped onto a separate corresponding visual channel. In case of combination of the same visual channels, those have been shown in two separate spaces.

ABSTRACT

In the information visualization reference model, *visual mapping* is the most crucial step in producing a visualization from a data set. The conventional visual mapping maps each data attribute onto a *single* visual channel (e.g., the year of production of a car to the position on the horizontal axis). In this work, we investigate *composite visual mapping*: mapping *single* data attributes onto *several* visual channels, each one representing one aspect of the data attribute (e.g., its order of magnitude, or its trend component). We first propose a table which allows us to explore the design space of composite mappings by offering a systematic overview of channel combinations. We expect that using more than one visual channel for communicating a data attribute increases the bandwidth of information presentation by displaying separable information on different aspects of data. In order to evaluate this point, we compare *horizon graph*, an existing technique which successfully adopts a composite visual mapping, with a selection of alternative composite mappings. We show that some of those mappings perform as well as –and in some cases even better than– horizon graph in terms of accuracy and speed. Our results confirm that the benefits of composite visual mapping are not limited to horizon graph. We thus recommend the use of composite visual mapping when users are simultaneously interested in several aspects of data attributes.

Index Terms: Information visualization, time series.

1 INTRODUCTION

Time series consist in data attributes changing through time, or events occurred along with the temporal axis. This association of data attributes’ values and states and the temporal information make time series one of the most common forms of recorded data in scientific (e.g., temperature variations), industrial (e.g., sensor data), financial (e.g., currency exchange rates) and many other applications.

Understanding temporal relationships in data allows us to learn from the past to predict and plan for the future [2]. Our excellent visual perception and reasoning combined with computers’ power and human-computer interactions make visualization of time series a valuable tool for understanding those relationships and patterns and eventually for problem solving and decision making [1].

It has always been a challenge to view all data and information in a limited screen space and this challenge becomes more present with the ever increasing volumes of datasets. Often, displaying every data variable and data record result in overplotting and visual clutter issues. In such situations, it becomes overwhelmingly impossible to obtain a clear view of the behaviour of data by applying the usual visualization techniques [3]. There are basically two options for addressing the such visual problems: (1) to artificially reduce the number of data points by means of sampling or aggregation, and (2) to enhance the drawing method to make it more space efficient than standard techniques [14]. The first option introduces involuntary information loss due to omission of data records, while the second may lead to complex visualizations that are not familiar to users.

Introducing a new visualization paradigm is not an easy endeavor, since it is very difficult to consider all data, visual, and task aspects for visualizing temporal data [2]. Time itself, for example, has many theoretical and practical aspects. For instance, it matters whether the user interprets time as a linearly ordered set of events or as a circularly recurring set of events. The number of data attributes to be analyzed at once, perceptual questions, and many more data and task related questions have to be considered as well.

In this work, we introduce a visual design space which aims to reduce information loss or increase the information communicated on different aspect of data by means of improving the visual mapping. We investigate an alternative method to visual mapping namely *composite visual mapping* while respecting common constraints (e.g., visual clutter and limited number of pixels). We carry out a systematic overview of the entire design space of composite mappings by means of an interactive table (Figure 1) and we highlight the existing techniques which adopt this approach to visual mapping. We compare a selection of our composite mappings with the state of the art in order to evaluate the general effectiveness of this approach. Based on our findings, we suggest the use of *composite visual mapping* for improving space-efficient visualizations of time series.

*e-mail: ali.jabbari@univ-grenoble-alpes.fr

†e-mail: renaud.blanch@univ-grenoble-alpes.fr

‡e-mail: sophie.dupuy@univ-grenoble-alpes.fr

The structure of this paper is as follows: Section 2 explains composite visual mapping and its theoretical basis; Section 3 highlights existing visualization techniques that benefit from composite visual mapping; Section 4 presents our systematic overview of composite visual mapping’s design space; Section 5 describes our evaluation of a selection of the new mappings and the results; and finally Section 6 puts forward the limitations and potentials of the current work.

2 BACKGROUND

The background of this work lies in the notion of visual channels as defined by Bertin [4]. The original channels as defined by Bertin in *Graphical Semiology* consist of: position (changes in the x, y location), size (change in length or area), shape (infinite number of forms), value (changes from light to dark), color (changes in hue at a fixed value), orientation (changes in alignment, angles), and texture (variation in grain or pattern). This list was later expanded by several publications and most notably by Mackinley [16] and Chen et al. [13] to include more than 30 visual channels. Most of the added channels are groupings or components of Bertin’s channels (e.g., size is divided into area, length and volume), while a few distinct channels such as motion (with the introduction of animation) and color saturation have also been added. Furthermore, Bertin defines characteristics for visual channels which help to determine if a visual channel is suitable for representing a given aspect of data. The five visual channel characteristics, according to Bertin [8], are:

- **Selective:** If a mark change in this channel can be perceived selectively in presence of other marks and channels,
- **Associative:** If marks can be grouped as a whole according to this channel,
- **Quantitative:** If a numerical reading of the channel is possible,
- **Order:** If changes in channel are perceived as ordered,
- **Length:** How many separate values can be perceived using this channel.

The choice of a proper visual channels for each data attribute is made based on these criteria with regards to the characteristics of the data and the tasks and goals of visual analysis.

The visualization pipeline (Figure 2) formalized by Card et al. [7] describes how a visualization is produced from raw data. After deriving focus data attributes in the *data transformation* step, they are linked to visual structures in the *visual mapping* step. This step bridges the gap between abstract data and forms perceivable by humans. Visual mapping is therefore the most critical step for ensuring expressiveness and effectiveness of the resulting visualization.

In conventional visual mapping, each data attribute is associated to a separate visual channel (e.g., car production of each year mapped on the height of bars in bar charts). However, dedicating a single

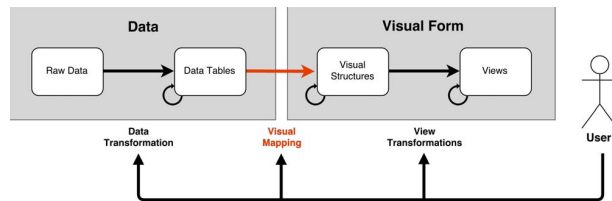


Figure 2: Visualization pipeline as formalized by Card et al. [7]. The step of visual mapping is the focus of this work.

visual channel to a data attribute may have some shortcomings. For instance, varying order of data magnitude can impact the efficiency of the visualization as some ranges of value for the visual channel remains unused and smaller magnitudes can be neglected in face of magnitudes several times larger. Other example of such limits is the visual clutter introduced in the visualization in presence of considerable noise in data. In this case, noisy variations of data that are not in the main interest of the user can obscure information and harm the accuracy of the user in reading the visual channel.

Selective attention theory asserts our ability to attend to one visual stimulus in an image (e.g., size of a symbol), while ignoring all other stimuli (e.g., color of the symbol) [19]. Evidence from psychological studies suggests that various combinations of visual channels may facilitate or inhibit selective attention [18]. These studies have further divided visual channels into *separable* (those that can be concentrated on independently) and *configural* (those that are highly interdependent). Although Bertin presented some ideas about combinations of visual channels, he did little research to empirically verify his hypotheses. Meanwhile, studies in cartography have shown that in dealing with map symbols composed of two separable visual channels, people can effectively perceive the information mapped on one visual channel while successfully filtering the information mapped on the other channel [20].

Based on these findings in psychology and cartography, in this work we support the idea of a form of redundant visual mapping for time-series. We think that using several separable visual channels for a single data attribute, namely a *composite visual mapping*, can improve the visualization pipeline for multiple reasons. Using several visual channels simultaneously, allows to distribute the information load to several channels, separate different facets of data (e.g., order of magnitude, high frequencies) and dedicate a full bandwidth visual channel to each of them. We think the main advantage of this approach is that users can focus more easily on the desired property of data without losing the rest of information. This approach allows user to separately analyze two or more aspects of data and perform different analytical tasks using each channel (e.g., trend detection by using the channel dedicated to low frequencies, and uncertainty assessment by using the channel dedicated to the noise component). Also, by distributing the information load on several visual channels, each channel becomes less congested and thus, more readable.

The importance of visual mapping and choice of visual channels has convinced us to concentrate our efforts on this step of the visualization pipeline. The potentials of combinations of visual channels promise improved visual forms that better represent data. However, depending on the nature of data and the analytical objectives of the user, various combinations of visual channels are imaginable. While some combinations have already been implemented and evaluated in visualization techniques, many others remain unexplored. This is the reason we think *composite visual mapping* is an approach worth a systematic study.

3 EXISTING COMPOSITE VISUAL MAPPINGS

Despite the large number of techniques for visualization of time series, only few of them use composite visual mapping as their approach to visual mapping.

3.1 Horizon Graph

Horizon graph [21] is one of the best known visualization techniques that implements a composite visual mapping. Horizon graph reduces space by dividing the chart into several bands and superposing them onto each other to create a layered form. Horizon graph uses the modulo transformation $v = \text{sgn}(v) \cdot (a \times M + b)$ to divide each data attribute to sign of the data attribute’s value (negative or positive), quotient a and remainder b . Each resulting components is then respectively mapped onto hue, saturation, and Y axis (Figure 3).



Figure 3: In horizon graph, the resulting quotients and reminders are mapped onto saturation and the vertical position, respectively. The sign of data attribute's value is mapped on hue (red for negative, blue for positive). Adapted from [15].

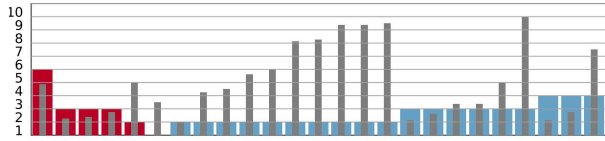


Figure 4: In Order of Magnitude Markers each data attribute is divided into a coefficient and a exponent of the logarithmic expression. Exponents are represented by colored bars and coefficients by grey ones. The sign of each data value is mapped onto hue (red/blue). Adapted from [5].

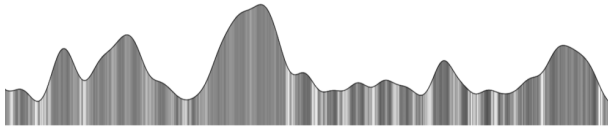


Figure 5: Slick Graph applies a frequency separation transformation on each data attribute. The low frequencies are mapped onto the vertical position (the smoothed line) and the high frequencies are encoded on the brightness of the area under the line. Adapted from [11].

The increase in data density enables the visualization system to display more data in a limited space. However, the deformations may obscure some patterns in data and the mental unstacking of layered charts may involve cognitive overload in some cases.

3.2 Order of Magnitude Markers

Order of Magnitude Markers (OOMM) [5] is a technique for representing numeral data that have varying orders of magnitude. The technique is aimed to allow the viewing of an entire data set, regardless of its varying order of magnitude in a limited space.

OOMM uses logarithmic transformation $v = \text{sgn}(v) \cdot (b \times 10^a)$ to divide each data attribute into sign of the data attribute's value (negative or positive), coefficient b , and exponent a . Each of these components are then respectively mapped onto hue, height of the coefficient bar, and height of the exponent bar. The two bars are distinguished by their color saturation (Figure 4).

Although Order of Magnitude Markers facilitates representation of discrete numerical data in the form of bar charts, it is not suitable for visualization larger temporal data due to its modest space management.

3.3 Slick Graph

Slick Graphs is a smoothing technique which aims to make time series visualizations more readable without losing information [11]. In other smoothing techniques [6], high frequencies' information is lost in the process of smoothing. Slick graphs on the other hand, present the whole information to the user while preserving the benefits of the

smoothing process. Compared to traditional smoothing techniques, Slick Graph has shown its superiority in various comparison tasks.

Slick graphs uses a frequency separation function to divide data attributes into high frequency and low frequency components. Low frequency components are mapped onto the vertical position (the line) and high frequency components are mapped onto the brightness (color value) of the area under the line (Figure 5).

4 COMPOSITE VISUAL MAPPING TABLE

In order to obtain a systematic overview of possible composite mappings, we created an interactive table which allows for exploration of different possible combinations of visual channels (Figure 1). In all charts of this table, the horizontal axis is reserved to time. We opted for linear representations of time because of the general familiarity of visualizations that use a 2D design with the horizontal axis corresponding to the temporal axis. However, we can integrate the same visual mappings into circular [12] or helix [22] representations of time series.

Viewer of the table can modify data parameters (e.g., data transformation function) and visual parameters (e.g., baseline saturation levels) in order to generate visualizations with composite mappings for a test data set. We use this table as a tool to present and explore the whole design space of composite visual mapping and its evaluation.

Our selection of visual channels is based upon the original list of primary channels as defined by Bertin with some modifications that follows. Our study focuses exclusively on abstract visualization of time series, and thus, we abandoned glyph related channels (e.g., shape). We included color saturation to this list as an independent channel missing in the original list. Dynamic visualization is not in the scope of this study, therefore we did not include motion (animation) in our channels. We use the following visual channels in the rest of this study:

- **Position** is defined as a mark's position along the Y axis. We reserved the X axis for time dimension to keep visualizations familiar to existing techniques;
- **Hue** is the degree to which a stimulus can be described as similar to or different from stimuli that are described as red, green, blue, and yellow;
- **Saturation** is the attribute related to chromatic intensity, ranging from black and white to fully colourful;
- **Value**, also known as *lightness*, is a representation of variation in perceived brightness of a color;
- **Texture** is the grain or pattern used in filling an area; and
- **Size** is the 2D (size) or 1D (length, thickness) of a mark (line or point).

Visual combination of two given channels is not unique. For instance, the texture channel used in Table 1 consists of vertical lines with varying thicknesses and yet other textures (e.g., dots or glyphs) can be used and affect the outcomes of the visual analysis. Also, the visual channels can be combined in different ways [9]. Most of the combinations in Table 1 use a shared space design for space-efficiency reasons. However, other forms (e.g., spatially separated) can also be applied. Spatial separation is however used when combining the same visual channels for both components (the diagonal of the table). Two bands are then placed side by side.

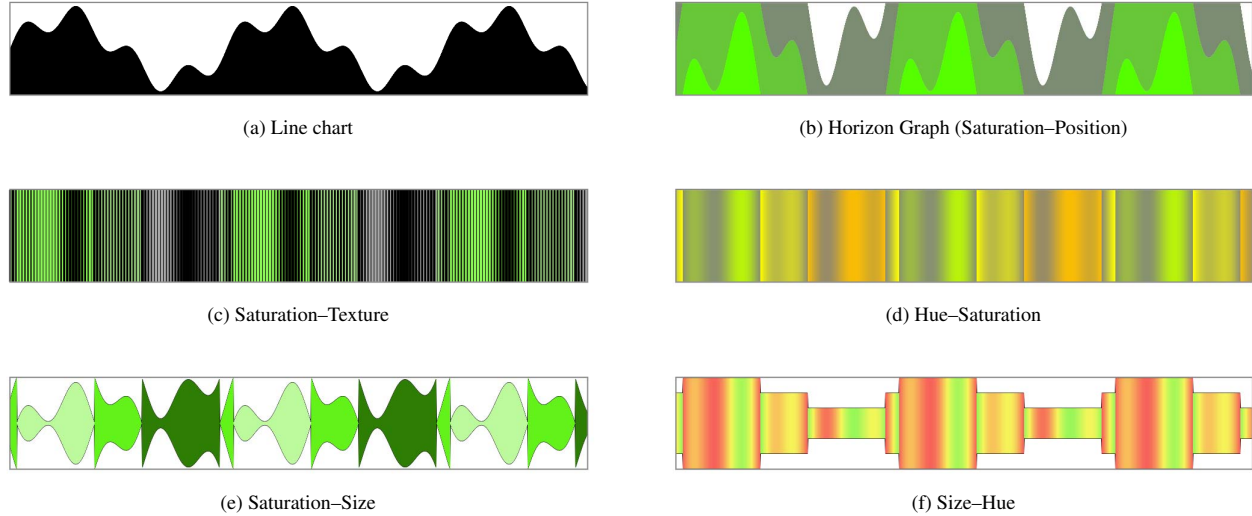


Figure 6: Complete selection of visual mappings examined in this study displaying a dummy data set. Except line chart which was included as the reference point, all other techniques use composite visual mapping with modulo division as the transformation function.

In the present version of the Table, we have studied two data transformations which are related to the order of magnitude of data attributes:

- **Modulo transformation:** Using modulo transformation, $v = a \times M + b$, each data attribute v is divided to two sub-attributes, quotient a and remainder b . In each cell of the table, the component a is mapped on the row's attributed visual channel and component b is mapped on the column's visual channel. Horizon graph uses modulo transformation as its transformation function where the quotient and remainder issued from the modulo transformation are mapped on saturation and position on Y-axis, respectively; and
- **Logarithmic transformation:** In logarithmic transformation, $v = b \times K^a$ (K being a constant calculated for each data set), each data attribute v is broken down to two sub-attributes, exponent a and coefficient b . The composite mapping table is thus constructed in the same way as with modulo transformation. In both transformations, component a communicates order of magnitude of the data attribute and component b communicates small variations of data attributes. Order of Magnitude Markers uses logarithmic transformation as its transformation function and is a redundant mapping of the resulting a and b onto Y-axis position.

The range of data values and the selected M or K determine the number possible discrete values for the sub-attribute a . When modulo transformation is used, the number of possible quotients of the division corresponds to number of layers or bands (as in horizon graphs). The list of transformations can be further extended to include other transformations (e.g., frequency separation).

From this table, we can systematically view the whole design space of composite visual mappings that can be constructed by combination of visual channels. As mentioned earlier, the presented designs are not the only possible forms of combined visual channels and can be modified and further enhanced. among the combinations introduced in this table, only few have been implemented in the literature. Horizon graph is the best known design which adopts composite visual mapping by means of modulo transformation. Since

earlier studies [15] have shown the effectiveness of horizon graphs in discrimination and value estimation tasks, we have decided to compare it with a selection of new composite mappings from our table to evaluate the effect of such visual mappings on those tasks.

5 APPROACH AND METHODS

Our goal in the following empirical study was to report the overall effectiveness of composite mappings before more rigorous evaluations of refined selected designs. We conducted an experiment to answer these questions:

1. **How accurate are the other new visual mappings?** How accurate and fast do participants perform in a discrimination and estimation task comparing to horizon graph?
2. **How does the choice of data transformation function affect users' performance in those tasks?** Does logarithmic transformation help evaluation? What is its impact on estimation speed?
3. **What is the impact of the number of layers?** We already know that increasing the number of layers will affect accuracy and estimation speed in horizon graphs. Is it the case with other composite visual mappings?

We were obligated to limit the number of mappings that we test to five (horizon graph and 4 new mappings) otherwise evaluation of all conditions would become practically impossible. We considered two groups of visual channels in selecting our test combinations: geometric channels (i.e. position and size), and retinal channels (i.e. hue, saturation, texture). The visual channels in the first group depend mostly on the physical properties of the marks (i.e. their physical size and position on the screen) and those in the second group mostly rely on perceptual capacities of human eye (e.g., perceived color). We tried to have a fair selection of combinations with channels from each of these groups for each sub-attributes. Our selection consisted of the following five composite visual mappings: Saturation-Position (horizon graph), Saturation-Texture, Hue-Saturation, Saturation-Size, and Size-Hue (Figure 6). Here and in the rest of this paper, the compound name of [Visual Channel 1]-[Visual Channel 2] points to the

combination at row [Visual Channel 1] and column [Visual Channel 2] of the composite visual mapping table with sub-attributes a and b , respectively.

Since we aimed to compare new mapping alternatives with the existing horizon graph, we decided to reproduce its experimental protocol [15]. Hence, participants of the experiment performed a discrimination and estimation task. We explain the task in Section 5.2.

In order to evaluate the effect of transformation function, one block of the trials (half of the total trials) had modulo transformation as their transformation function, while the other block had logarithmic transformation. We divided the subjects into two groups and we permuted the order of the blocks for the two groups to reduce the learning effect.

In order to observe the impacts of layering, we tested 2 and 3 layers variations of each mapping. We used the same configuration for 2 different data transformations: Modulo transformation and logarithmic transformation. This resulted in $5(\text{mapping}) \times 2(\text{layers}) \times 2(\text{transformation})$ conditions. We generated 5 trials for each condition, and thus, our experimental design consisted of a total of $5 \times 2 \times 2 \times 5 = 100$ trials per participant.

5.1 Chart Generation

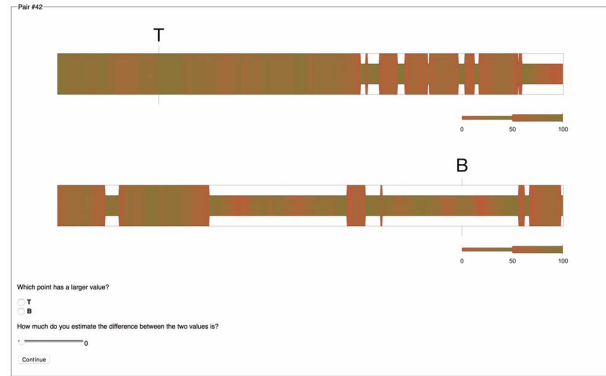
For each chart, we generated a time series using a moving average smoothing over a random walk. Time series' minima and maxima were then translated onto 0 and 100 respectively. This resulted in having the same value range for all charts. The charts were made on HTML 5 Canvas elements using JavaScript. All charts had a width of 2000 pixels and a height of 160 pixels. All charts measured 233×19 mm on our screen. In an attempt to preserve perceptual uniformity, we used CIELUV color space.

5.2 Experimental Procedure

We recruited 12 participants among people working at our laboratory. The participants, 10 male and 2 female, aged between 22 and 28 and were all computer science students or degree-holders. For each participant, the experiment begins with a tutorial on each of the selected mappings and the global procedure of the experiment. Participant, then, takes control of the experimental interface and starts with filling a form with information on her age, vision and education. After this step, the first trial appears and the main experiment begins.

Our experimental interface was entirely created with JavaScript. In each trial, the participant is presented with a pair of two separate charts, one positioned below the other one. The two charts of each pair had identical visual configurations, while conveying two distinct time series. Each chart of a pair was marked either T or B in a fixed position through all trials (Figure 7). At each trial, the participant was asked to report whether position T or B represents a greater value via radio buttons marked T and B . Participants were then asked to estimate the absolute value difference between the two positions. They responded by using a slider without tick marks in order to avoid anchoring effects [17]. The value selected via the slider was shown as numerals next to the slider. Upon clicking on "Next" button, the participant's responses, as well as the completion time for the trial was recorded and a new pair of charts replaced the former ones. The correct answer for the trial was not communicated to the participant to avoid learning effects. Once an answer was submitted, participants could not go back to review their answers.

The experiment consisted in 100 trials divided into 10 blocks of 10 trials. The trials in each of the 10 blocks had the same mapping and data transformation. Participants could voluntarily take breaks between blocks and resume when they were ready. The experiment was run in Google Chrome web browser version 59.0 on a 27-inch Apple iMac with Retina 5K display in full brightness in a room with constant ceiling lighting and no sunlight. During an interview at the end of the experiment, participants shared their views on the experiment and different visualizations with the experimenter.



Which point has a larger value?

- T
 B

How much do you estimate the difference between the two values is?



Continue

Figure 7: Interface of the experiment, presenting a pair of charts with Size-Hue mapping with three-layers modulo transformation. (top) global view, (bottom) detail on the questions and answers.

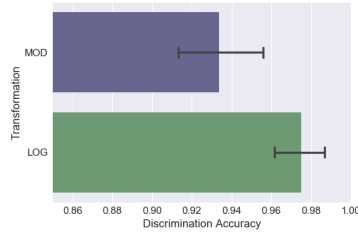
5.3 Results

For the analysis, we removed an outlier trial which received wrong answer from all of the participants. This should be due to the extremely small value difference between the two markers (0.05 in a scale of 0 to 100). However, it can be a basis for further study of equality cases in order to evaluate users' estimation behavior. Otherwise, none of the participants demonstrated wildly inaccurate performance in any block of the experiment. The participants were divided into two groups: the first group started the experiment with modulo transformations and the second group began with logarithmic transformations. The two groups demonstrated very similar accuracy levels (closely in the same intervals of confidence) which rejects the possibility of a significant learning effect. In accordance with other studies [10], we reported logarithmic absolute error as a measure of estimation error.

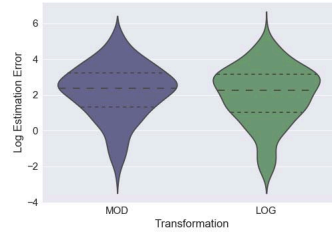
5.3.1 Performance by Data Transformation

For both modulo and logarithmic transformations, discrimination accuracy averaged above 90%. However, participants performed slightly better with logarithmic transformation in the discrimination task (Figure 8a). On the other hand, the estimation error does not differ between the two transformations. Figure 8b shows the distribution of logarithmic absolute error for the two transformations.

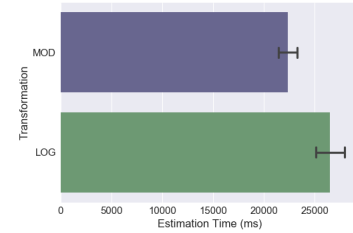
In generating the visualization pairs with logarithmic transformation, we calculated the base depending on the minimum and maximum of each data set in order to ensure having the chosen number of "layers". This varying base of logarithm results in varying value ranges on those charts' legends which may require extra mental calculation to estimate value differences. Participants reported this point in the follow up interviews: they needed more mental calculations to deduce the value difference. Indeed, our results show longer completion times for logarithmic transformation (Figure 8c).



(a) Average discrimination accuracy for modulo and logarithmic transformations. Black brackets show the 95% confidence intervals.

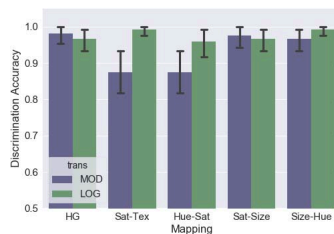


(b) The distribution of logarithmic absolute error for modulo and logarithmic transformations. They share similar distribution of error. The dashed lines represent the median, the first, and the third quartiles.

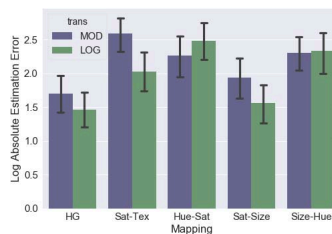


(c) Completion time for modulo and logarithmic transformations. Logarithmic transformation extends the completion time due to the irregular values it introduces in chart legends.

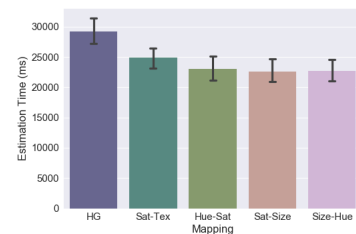
Figure 8: Effect of data transformation function: modulo vs. logarithmic transformation.



(a) Discrimination rates for different mappings. Sat-TeX with modulo transformation performed the worst. The effect was not observed with logarithmic transformation. Black brackets show the 95% confidence intervals



(b) Lower absolute errors for Horizon Graph, while the difference (in units) is not significant in this scale.



(c) Completion time for the five visual mappings: Horizon Graph performed the slowest, while others had the same levels of completion time.

Figure 9: Effect of visual mapping: Horizon Graph (saturation-position) vs. our selection of composite mappings.

5.3.2 Performance by Visual Mapping

Figure 9a shows discrimination accuracy rates for the five selected visual mappings and respective rates for the two data transformations. Sat-TeX and Hue-Sat mappings with modulo transformations had the lowest average discrimination accuracy, while other conditions had similar rates all above 90%. Interestingly, in the former two mappings, charts with logarithmic transformation had much higher discrimination accuracy rates above 90%.

Absolute estimation errors (Figure 9b) were lower for Horizon Graph and Sat-Size mapping (at around 5 units) comparing to the others (at around 8 units). The difference, however, does not seem significant in a scale of 100 units.

The results show longer completion times for Horizon Graph (Figure 9c). The other mappings exhibit lower levels that are similar.

5.3.3 Performance by Number of Layers

The design of the experiment allowed for either two or three discrete values for the sub attribute related to the order of magnitude (i.e. quotient for modulo transformation, power coefficient for logarithmic transformation). This range of values results in two or three levels (layers or bands) of the attributed visual channels. We limited the number of layers to three because earlier studies [15] have shown that higher number of layers significantly increase the error.

The results show that a three-layers design compared to a two-layers design leads to higher discrimination accuracy rates, lower estimation errors and similar completion times (Figure 10).

5.3.4 Performance by Real Value Differences

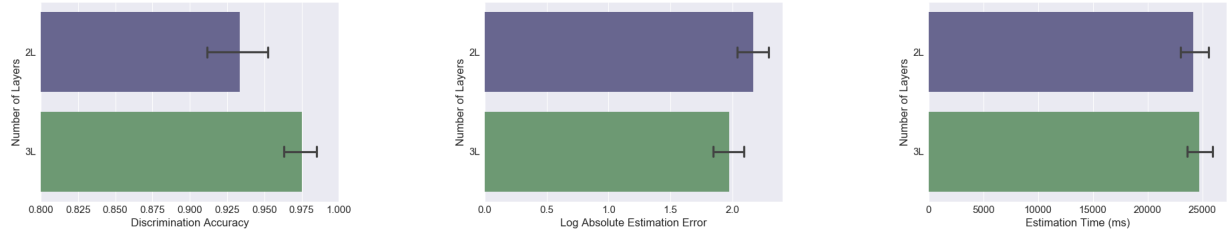
Figure 11 shows the effect on discrimination accuracy for the two transformations. Modulo transformation has lower scores in very small and very large actual differences. On the other hand, logarithmic transformation score retains the same levels of scores across ranges of actual difference.

Figure 12 reports the effect on estimation error for the five mappings. Overall, larger actual differences between the values result in higher estimations errors. However, horizon graph and Saturation–Size demonstrate similar behavior and are less prone to the actual value difference.

Figure 13 shows that the estimation time increases across the board. Horizon Graph is the mapping the most affected by this, while Size-Hue is the least susceptible.

5.4 Discussion

Our results confirm our first hypothesis that the accuracy and speed of other alternative composite mappings can attain levels similar to horizon graphs for the same discrimination and estimation task. The discrimination accuracy rates averaged above 90% for all of the mappings. Mean estimation errors were similar across different mappings, while they were slightly lower for horizon graphs and Sat-Size. This can be explained by the fact that size, and similarly position which in case of horizon graph results in larger areas, are the most easily perceived visual channels for most users [10]. In matters of speed, horizon graphs showed to be the slowest mapping. During the interviews with the participants following the experiment, some



(a) Discrimination accuracy: better correct discrimination with three-layers design

(b) Logarithmic error: lower estimation error with three-layers design

(c) Completion time: no meaningful difference between the two- and three-layers design

Figure 10: Effect of layering: two-layers vs. three-layers design. Higher layering was excluded due to the proven lower accuracy according to earlier studies

noted that they found the mental unfolding of the horizon graphs to be a cognitively demanding task. Although this can partially explain the slower performance for horizon graphs, more training may reduce the effect. In addition, higher familiarity and confidence of users in geometric visualizations including horizon graphs might drive them to spend more time in hope of achieving more accurate estimations. With current evaluation, we cannot identify the cause of longer estimation times in horizon graphs.

Based on the interested aspects of data, many transformation functions can be imagined for dividing data attributes. In our study we focused on modulo and logarithmic transformations, two functions that are more related to the magnitude of data and that have already been used in the existing horizon graphs and order of magnitude markers. Discrimination accuracy averaged slightly higher for logarithmic transformation. This is explained by the higher discrimination accuracy averages for very small and very large differences (Figure 11). This effect may be explained by the fact that with logarithmic transformation, two points both with small values are more probable to fall into the different regions of the visual channel representing the magnitude component. The use of logarithmic transformation did not impact the estimation error, but decreased the speed of users. Overall, the effect of transformation function was limited.

Consistent with previous studies [15], using more “layers” which translates to more discrete values for the magnitude component helps users by increasing their discrimination accuracy and lowering estimation error without sacrificing the speed. However, as the previous studies have shown, more than three layers can harm the performance of the user in estimation task and thus, should be avoided.

We found that the real difference of values between the two points has an impact on participants’ estimation accuracy (Figure 12).

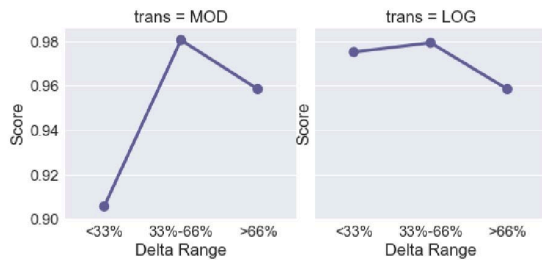


Figure 11: Effect of the real value difference on discrimination accuracy: Modulo vs. logarithmic transformation.

All five mappings show bigger errors with larger real value differences, but the effect is less prominent for horizon graphs and Size-Saturation. We observe that while in small real value differences all mappings show similar error levels, those two mappings show lower errors at large real value differences.

Also, the increase in real value difference has negative impact on participants’ speed, with horizon graph being more prone to this effect in this regard (Figure 13).

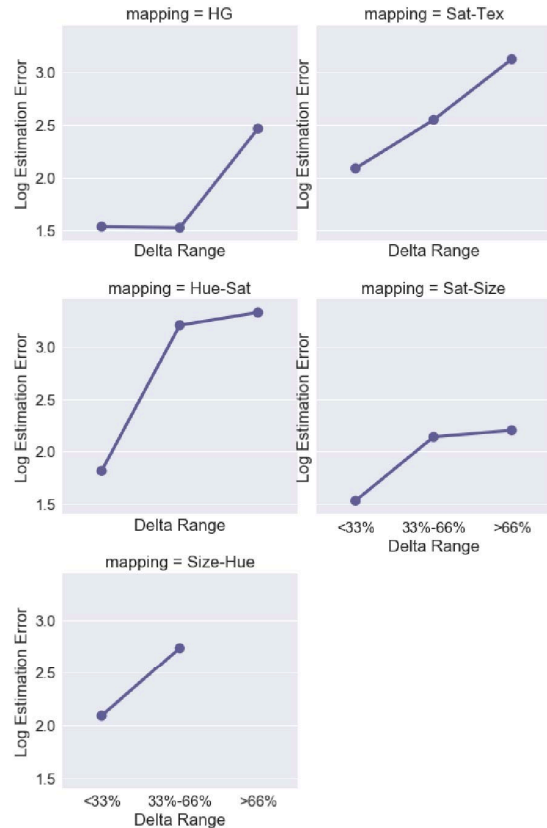


Figure 12: Effect of the real value difference on logarithmic error: comparing the five mappings.

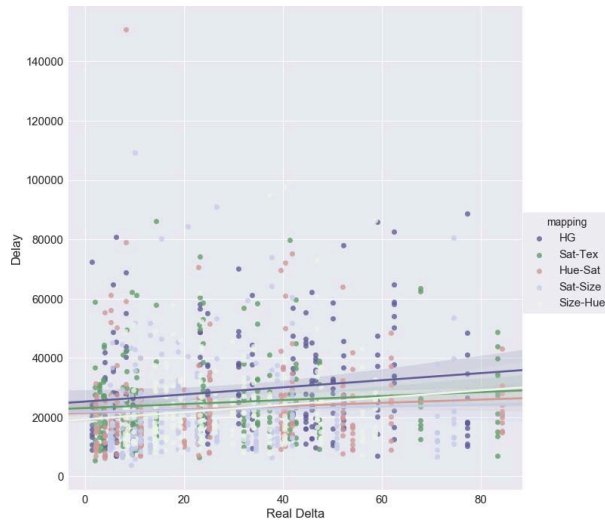


Figure 13: Effect of the real value difference on completion time: linear regression lines with 95% confidence intervals. Horizon graph (HG) is the most vulnerable to the real value difference.

6 LIMITATIONS AND FUTURE WORK

In this work, we examined a selection of composite mappings among the possible mappings that we presented in our design space table. Our aim was to show the general benefits of our approach to visual mapping in discrimination and estimation tasks. The empirical study allowed us to report some observed trends, yet more comprehensive evaluations are clearly needed for statistically significant results. Further studies may discover advantages of particular mappings in the appropriate context with improved configuration.

One limitation of the present study is that we only measured the performance of users in a narrow task of discrimination and estimation. While this is a common task in visual analysis of time series, the effects of composite mapping on other tasks such as trend perception and graphical perception of rates of change remain to be investigated. Another limitation of this work is that for each combination of visual channels, several visual designs are imaginable, and we only tested one particular design. Therefore, design improvements may enhance the performance of the technique.

In our evaluation, we only tested one type of data. While smoothed random walks are relevant data sets for discrimination and estimation tasks, other types of data sets (e.g., noisy data) are suitable for other analytical tasks (e.g., trend detection). Such cases require future evaluations.

Although horizon graphs were conceived as a means to reduce the space, earlier experiments [15] have discovered that estimation error is higher for smaller charts. We suspect that this effect is due to the dependence of horizon graphs on position channel and hence, the size of the chart. As a future step of our study, we are interested in looking at the effect of the size of the chart in estimation accuracy with different composite mappings. We suspect that the composite mappings which do not use position are less vulnerable to the effect of chart size.

7 CONCLUSION

In this work we have presented new visual designs for visualization of time series with the goal of increasing information presentation bandwidth. Our approach to the visual mapping results in a large design space of mappings, from which, only few have been examined in the literature. We introduced one visual design for each

combination, only to illustrate the possible design space for time series visualization. Our experiment has shown the state-of-the-art accuracy of a selection of these mappings in a discrimination and estimation task. During interviews with participants of our evaluation, they found the new mappings straightforward to learn.

We found composite visual mapping an interesting basis for new time series visualizations. Future refined visual designs can be created based on this approach to visual mapping. Compactness and vast visual variability of composite mappings make them promising for many contexts and analytical tasks. It is of our great interest to further the investigation of composite visual mappings in terms of sensibility to chart size, learnability, and measures other than value estimation accuracy.

ACKNOWLEDGMENTS

The authors would like to thank all the people involved in the experiment. This work has been funded by the French program Nano 2017.

REFERENCES

- [1] W. Aigner. *Interactive Visualization and Data Analysis : Visual Analytics With a Focus on Time*, 2013. Habil. thesis, Technische Universität Wien.
- [2] W. Aigner, S. Miksch, W. Müller, H. Schumann, and C. Tominski. Visualizing time-oriented data – A systematic view. *Computers and Graphics*, 31(3):401–409, 2007. doi: 10.1016/j.cag.2007.01.030
- [3] N. Andrienko and G. Andrienko. *Exploratory Analysis of Spatial and Temporal Data*. Springer, 2006. doi: 10.1007/3-540-31190-4
- [4] J. Bertin. *Sémiologie Graphique. Les diagrammes, les réseaux, les cartes*. Gauthier-Villars, 1967.
- [5] R. Borgo, J. Dearden, and M. W. Jones. Order of magnitude markers: An empirical study on large magnitude number detection. *IEEE TVCG (Proc. InfoVis 2014)*, 20(12):2261–2270, 2014. doi: 10.1109/TVCG.2014.2346428
- [6] L. Byron and M. Wattenberg. Stacked graphs – geometry & aesthetics. *IEEE TVCG (Proc. InfoVis 2008)*, 14(6):1245–1252, 2008. doi: 10.1109/TVCG.2008.166
- [7] S. Card and J. Mackinlay. The structure of the information visualization design space. In *Proc. IEEE InfoVis'97*, pp. 92–99, 1997. doi: 10.1109/INFVIS.1997.636792
- [8] S. Carpendale. *Considering Visual Variables as a Basis for Information Visualisation*. Technical report, Department of Computer science, University of Calgary, 2001.
- [9] M. Chen, S. Walton, K. Berger, J. Thiyyagalingam, B. Duffy, H. Fang, C. Holloway, and A. E. Trefethen. Visual multiplexing. *Computer Graphics Forum (Proc. EuroVis 2014)*, 33(3):241–250, 2014. doi: 10.1111/cgf.12380
- [10] W. S. Cleveland and R. McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *J. Amer. Statist. Assoc.*, 79(387):531–554, 1984. doi: 10.1080/01621459.1984.10478080
- [11] R. Dautriche, R. Blanch, A. Termier, and M. Santana. TraceViz: A visualization framework for interactive analysis of execution traces. In *Proc. IHM 2016*, pp. 115–125, 2016. doi: 10.1145/3004107.3004127
- [12] J. Fuchs, F. Fischer, F. Mansmann, E. Bertini, and P. Isenberg. Evaluation of alternative glyph designs for time series data in a small multiple setting. In *Proc. ACM CHI 2013*, pp. 3237–3246. ACM, 2013. doi: 10.1145/2470654.2466443
- [13] C. Hansen, M. Chen, C. Johnson, A. Kaufman, and H. Hagen, eds. *Scientific visualization: uncertainty, multifield, biomedical, and scalable visualization*. Springer, 2014. doi: 10.1007/978-1-4471-6497-5
- [14] M. C. Hao, U. Dayal, D. A. Keim, and T. Schreck. Multi-resolution techniques for visual exploration of large time-series data. In *Proc. EuroVis 2007*, p. 8, 2007. doi: 10.2312/VisSym/EuroVis07/027-034
- [15] J. Heer, N. Kong, and M. Agrawala. Sizing the horizon: the effects of chart size and layering on the graphical perception of time series visualizations. In *Proc. ACM CHI'09*, pp. 1303–1312, 2009. doi: 10.1145/1518701.1518897

- [16] J. Mackinlay. Automating the design of graphical presentations of relational information. *ACM Trans. Graph.*, 5(2):110–141, 1986. doi: 10.1145/22949.22950
- [17] J. Matejka, M. Glueck, T. Grossman, and G. Fitzmaurice. The effect of visual appearance on the performance of continuous sliders and visual analogue scales. In *Proc. ACM CHI 2016*, pp. 5421–5432, 2016. doi: 10.1145/2858036.2858063
- [18] E. S. Nelson. Using selective attention theory to design bivariate point symbols. *Cartographic Perspectives*, (32):6–28, 1999. doi: 10.14714/CP32.625
- [19] E. S. Nelson. The impact of bivariate symbol design on task performance in a map setting. *Cartographica*, 37(4):61–78, 2002. doi: 10.3138/V743-K505-5510-66Q5
- [20] E. S. Nelson. Designing effective bivariate symbols: The influence of perceptual grouping processes. *Cartography and Geographic Information Science*, 27(4):261–278, 2010. doi: 10.1559/152304000783547786
- [21] T. Saito, H. N. Miyamura, M. Yamamoto, H. Saito, Y. Hoshiya, and T. Kaseda. Two-tone pseudo coloring: Compact visualization for one-dimensional data. In *Proc. IEEE InfoVis 2005*, pp. 173–180, 2005. doi: 10.1109/INFOVIS.2005.35
- [22] C. Tominski, P. Schulze-Wollgast, and H. Schumann. 3d information visualization for time dependent data on maps. In *Proc. IV'05*, pp. 175–181, 2005. doi: 10.1109/IV.2005.3